

Medições, erros aleatórios e o filtro de Kalman

Marco Costa, marco@ua.pt

*ESTGA-UA, Escola Superior de Tecnologia e Gestão de Águeda
CIDMA, Centro de Investigação e Desenvolvimento em Matemática e Aplicações
Universidade de Aveiro*

1. Introdução

As medições experimentais têm sempre uma incerteza associada. Essa incerteza traduz a impossibilidade de se obter um valor verdadeiro através de um procedimento de medição, em regra, com o auxílio de um ou mais instrumentos de medida que, indubitavelmente induzem erros nas medições. Esses erros podem ser aleatórios no sentido em que são inevitáveis e imprevisíveis, dependendo das condições instrumentais ou ambientais em que as medições são obtidas; ou não aleatórios, por exemplo se forem sistemáticos, sendo suscetíveis de serem corrigidos. São os erros aleatórios que suscitam a necessidade da aplicação de técnicas estatísticas para a sua quantificação ou redução.

A incerteza associada a uma medição única pode ser analisada quanto à sua precisão se considerarmos o erro máximo do instrumento utilizado obtendo-se, desta forma, um majorante para o erro absoluto. Ou, caso seja mais relevante, pode obter-se um majorante para o erro relativo associado a essa medição. Esta é a abordagem que, em regra, se adota no âmbito na Análise Numérica e no tratamento de observações ou de valores resultantes de operações matemáticas sobre medições.

No entanto, em muitas situações em Engenharia dispomos de medições que, além de terem associadas a incerteza inerente ao instrumento ou método de medição, e cujo erro pode ser eventualmente majorado, são resultantes de outros fatores aleatórios que, em regra, não se consegue controlar ou majorar. Por outro lado, em muitos casos práticos dispomos de medições sucessivas que podemos considerar como sendo observações de um processo estocástico, isto é, constituem uma série temporal, e que cuja variabilidade pode ser atribuída a fatores ambientais e/ou instrumentais no âmbito dos procedimentos de medição.

É nestes casos que a análise e a modelação de séries temporais na perspetiva da sua representação em espaço de estados pode ser uma mais valia no tratamento de medições num paradigma estocástico. A modelação de séries temporais, através de modelos de espaço de estados, tem-se evidenciado como uma abordagem bastante útil na análise de séries temporais, cujos principais objetivos sejam a previsão a curto prazo e/ou a obtenção de estimativas “filtradas”.

De facto, os modelos de espaço de estados têm uma estrutura *markoviana* que permite obter estimativas de variáveis não observáveis e previsões suscetíveis de serem atualizadas quando nova informação é recolhida por aplicação do conhecido filtro de Kalman (Kalman, 1960). O filtro de Kalman é um conjunto de equações recursivas que são estabelecidas para um modelo que admita uma representação de espaço de estados que permitem obter, em cada instante, estimativas e previsões e os respetivos erros quadráticos médios.

2. Modelo de espaço de estados e o filtro de Kalman

A variável de interesse X_t , que designamos por *estado*, e que não é observável, mas sobre a qual se pretende obter estimativas e, eventualmente, previsões, evolui ao longo do tempo segundo um modelo autorregressivo da forma

$$X_t = \phi X_{t-1} + \varepsilon_t \quad (1)$$

isto é, o estado no instante t , X_t , é uma correção do estado no instante anterior “contaminado” com um erro aleatório ε_t . O coeficiente ϕ é designado por coeficiente autorregressivo e o erro aleatório ε_t é assumido como sendo um ruído branco gaussiano, com $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ e $E(\varepsilon_t \varepsilon_s) = 0, \forall t \neq s$.

O estado X_t , não sendo observável, é indiretamente medido após a sua multiplicação por um fator h_t conhecido cujo resultado é “corrompido” por outro ruído branco gaussiano $e_t \sim N(0, \sigma_e^2)$, com $E(e_t e_s) = 0, \forall t \neq s$, cujo resultado observável é a variável Y_t , isto é,

$$Y_t = h_t X_t + e_t. \quad (2)$$

Em regra, os erros ε_t e e_t são assumidos como não sendo correlacionados, isto é, $E(\varepsilon_t e_s) = 0, \forall t, s$.

A equação (1) é a designada *equação de estado* e descreve a evolução temporal da variável de interesse, enquanto que, a equação (2) é a chamada *equação de observação* relacionando o estado não observável X_t com a variável de *medida* Y_t , esta observável. Os erros ε_t e e_t são comumente designados nas Engenharias apenas como *ruídos*. Um modelo que admita uma representação sobre a forma das equações (1)-(2) diz-se que admite uma *representação de espaço de estados*.

O modelo (1)-(2) é bastante versátil podendo ser interpretado de diversas formas e em vários contextos. Em particular, podemos entender o estado X_t como o verdadeiro valor da medição no instante t , sendo que a variável Y_t representa a medição efetivamente observada. Este modelo incorpora duas fontes de variabilidade, os dois ruídos brancos, onde ε_t pode representar a variabilidade inerente ao processo que se pretende analisar e o erro e_t pode representar a variabilidade induzida, por exemplo, pelo instrumento de medição e/ou condições ambientais. Noutros contextos, o modelo (1)-(2) pode ser interpretado como um modelo de regressão linear cujo declive não é um parâmetro (determinístico) sendo uma variável aleatória, permitindo uma evolução dinâmica ao longo do tempo.

Os modelos ARMA (Modelos Autorregressivos e de Médias Móveis) são uma classe de modelos que admitem uma representação se espaço de estados, permitindo, por exemplo a obtenção das estimativas de máxima verosimilhança dos parâmetros através desta abordagem. Existem várias extensões do modelo (1)-(2) que têm sido propostas na literatura. Por exemplo, o estado e a variável de medida podem ser vetores aleatórios com dimensões adequadas ou as equações de observação e de estado podem ter outras configurações tais como incluir outros coeficientes ou matrizes de design/planeamento.

O filtro de Kalman é um conjunto de equações recursivas que exprimem recursivamente, em cada instante t , os estimadores de erro quadrático mínimo sob a hipótese de normalidade dos ruídos ε_t e e_t , ou os melhores estimadores lineares centrados quando os

ruídos não são normais, uma que são as projeções ortogonais do estado sobre as variáveis observadas até ao instante t ; bem como os respetivos erros quadráticos médios das estimativas.

Assim, seja $X_{t|t-1}$ o estimador do estado X_t com a informação disponível até ao instante $t - 1$, isto é, baseado em $\mathcal{Y}_{t-1} = (Y_1, Y_2, \dots, Y_{t-1})$ e $P_{t|t-1}$ o respetivo erro quadrático médio. O preditor de Y_t , baseado em \mathcal{Y}_{t-1} , que denotamos por $Y_{t|t-1}$, resulta da linearidade da equação de medição e é dado por

$$Y_{t|t-1} = h_t X_{t|t-1}. \quad (3)$$

Mas quando no instante t temos disponível a medição Y_t podemos atualizar a última estimativa do estado através do erro de predição ou *inovação*, $\eta_t = Y_t - Y_{t|t-1}$, pela equação

$$X_{t|t} = X_{t|t-1} + k_t \eta_t \quad (4)$$

onde

$$k_t = P_{t|t-1} h_t (h_t^2 P_{t|t-1} + \sigma_e^2)^{-1} \quad (5)$$

é o designado *ganho de Kalman*. O estimador $X_{t|t}$ permite a obtenção da estimativa filtrada do estado com a informação disponível até ao instante t cujo erro quadrático médio é dado por

$$P_{t|t} = P_{t|t-1} (1 - k_t h_t). \quad (6)$$

No entanto, através da equação de estado, no instante t podemos obter a previsão do estado para o instante seguinte, $t + 1$, pela equação

$$X_{t+1|t} = \phi X_{t|t}$$

com erro quadrático médio dado por

$$P_{t+1|t} = \phi^2 P_{t|t} + \sigma_\varepsilon^2. \quad (7)$$

As equações recursivas são inicializadas com os valores $X_{1|0}$ e $P_{1|0}$ que podem decorrer do caso prático ou podem ser estimadas no procedimento da estimação dos restantes parâmetros que referiremos mais à frente. Os estimadores do filtro de Kalman são ótimos no sentido do menor erro quadrático médio quando os ruídos são normais, isto é,

$$X_{t|t} = E(X_t | Y_1, Y_2, \dots, Y_t) \text{ e } X_{t|t-1} = E(X_t | Y_1, Y_2, \dots, Y_{t-1}). \quad (8)$$

Sob a normalidade dos ruídos ε_t e e_t , os parâmetros desconhecidos $\Theta = (\phi, \sigma_\varepsilon^2, \sigma_e^2)$, entre outros se existirem, podem ser estimados pelo método da máxima verosimilhança. Nestas condições, a log-verosimilhança de uma amostra $\mathcal{Y}_t = (Y_1, Y_2, \dots, Y_n)$ pode ser escrita através das distribuições condicionais, obtendo-se

$$\ln L(\theta | \mathcal{Y}_t) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \ln(\omega_t) - \frac{1}{2} \sum_{t=1}^n \ln(\omega_t^{-1} \eta_t^2)$$

onde $\omega_t = h_t^2 P_{t|t-1} + \sigma_e^2$.

As estimativas de máxima verosimilhança, $\hat{\Theta}_{MV}$, são obtidas maximizando a log-verosimilhança através de métodos numéricos como o algoritmo EM (Dempster et al., 1977) ou o método de Newton-Raphson (Harvey, 1996), isto é,

$$\hat{\Theta}_{MV} = \underset{\theta}{\operatorname{argmax}} \ln L(\theta | y_t). \quad (9)$$

É possível obter-se estimativas para os coeficientes desconhecidos através de *packages* e de rotinas computacionais, como o *package* dlm do ambiente R (Petrís, 2010) que incorpora várias funções que permitem modelar e estimar os parâmetros dos modelos de espaço de estados usuais, bem como obter as estimativas e as previsões obtidas através do filtro de Kalman.

Sob algumas condições gerais, os estimadores da máxima verosimilhança são assintoticamente normais, tendo-se,

$$\sqrt{n}(\hat{\Theta}_{MV} - \Theta) \xrightarrow{D} N(0, I^{-1}(\Theta)), \quad (10)$$

onde $I(\Theta)$ é a matriz de informação, cujo elemento (i, j) pode ser aproximado por

$$I_{ij}(\Theta) \approx \frac{1}{2} \sum_{t=1}^n \frac{1}{\omega_t^2} \frac{\partial \omega_t}{\partial \theta_i} \frac{\partial \omega_t}{\partial \theta_j} + \sum_{t=1}^n \frac{1}{\omega_t} \frac{\partial \eta_t}{\partial \theta_i} \frac{\partial \eta_t}{\partial \theta_j},$$

e as derivadas parciais podem ser obtidas recursivamente aquando da implementação do filtro de Kalman ou através da diferenciação numérica.

3. Alguns comentários

As variâncias dos ruídos

As grandezas das variâncias das equações de observação e de estado, respetivamente σ_e^2 e σ_ε^2 , indicam a “confiabilidade” do sensor (isto num contexto mais tecnológico) relativamente à capacidade de estimar o estado.

Por exemplo, se $\sigma_e^2 \gg \sigma_\varepsilon^2$ significa que o erro de observação, em regra associado ao sensor, isto é, associado ao método ou procedimento de medição, é bastante superior ao error inerente ao fenómeno que se pretende estimar, o estado. Caso contrário, se $\sigma_e^2 \ll \sigma_\varepsilon^2$ o sensor diz-se bastante confiável no sentido em que o método de medição acrescenta pouca variabilidade ao estado, logo este será mais fácil de estimar (é observado com pouco ruído).

Por exemplo, na Figura 1 estão representadas duas realizações do modelo de um passeio aleatório mais um erro definido pelas equações

$$Y_t = X_t + e_t$$

$$X_t = X_{t-1} + \varepsilon_t$$

considerando $\sigma_e^2 = 10$ e $\sigma_\varepsilon^2 = 1$ e $\sigma_e^2 = 1$ e $\sigma_\varepsilon^2 = 5$, o estado e as estimativas do estado obtidas pelo filtro de Kalman.

Podemos verificar que quando σ_e^2 tem uma grandeza bastante maior que σ_ε^2 as estimativas do estado atualizadas pelo filtro de Kalman em cada instante, $X_{t|t}$, tendem a não acompanhar as observações de Y_t , com $t = 1, 2, \dots, n$, isto é, o filtro aplica a equação de estado, pouco atualizando a previsão $X_{t|t-1}$ quando Y_t é observada.

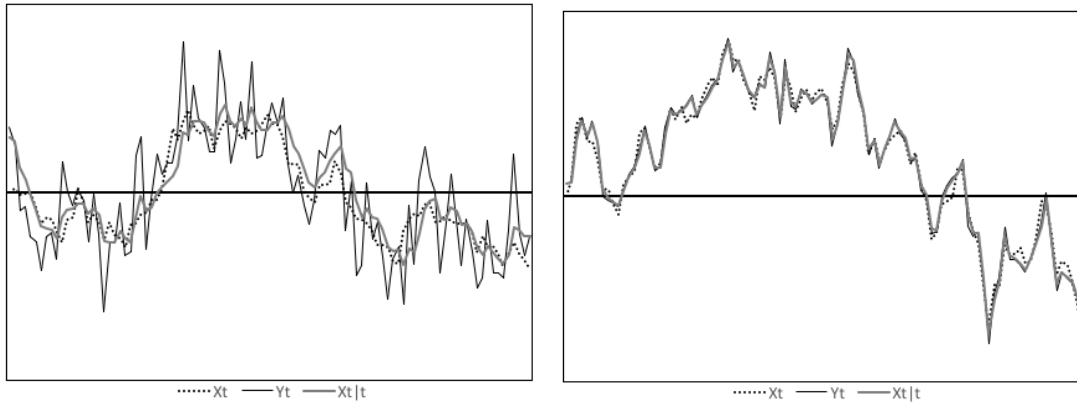


Figura 1: (à esquerda) uma realização de um processo do tipo $\sigma_\epsilon^2 \gg \sigma_x^2$; (à direita) uma realização de um processo do tipo $\sigma_\epsilon^2 \ll \sigma_x^2$.

Repare-se que, neste caso, quando $\sigma_\epsilon^2 \rightarrow +\infty$, temos que o ganho de Kalman tende para 0, $k_t \rightarrow 0$, pelo que a estimativa atualizada será semelhante à previsão, $X_{t|t} \approx X_{t|t-1}$.

Erros não gaussianos

O exemplo anterior ilustra a importância da estimação dos parâmetros de um modelo de espaço de estados, em particular as variâncias dos erros. Quando os erros são normais a estimação dos parâmetros desconhecidos é, em regra, realizada através do método da máxima verosimilhança (Eq. 9), como já discutido anteriormente.

No entanto, em muitas aplicações (ambientais, económicas, industriais, etc.) a normalidade dos erros nem sempre é garantida ou é rejeitada no procedimento de validação do modelo, nomeadamente através da análise da normalidade das inovações que, sob a normalidade dos erros, seguem a distribuição $\eta_t = Y_t - Y_{t|t-1} \sim N(0, \omega_t)$. Por outro lado, em modelo mais complexos (por exemplo, modelos de espaço de estados multivariados que aqui não são abordados por uma questão de simplicidade), nem sempre é garantida a unicidade de um máximo global da função de log-verosimilhança, $\ln L(\theta|y_t)$ (Hamilton, 1994; pág. 387).

Neste sentido, têm sido propostas abordagens alternativas como estimadores não-paramétricos (Alpuim, 1999; Costa & Alpuim, 2010; Gonçalves & Costa, 2013); ou a abordagem Bayesiana com recurso a técnicas de Monte Carlo via cadeias de Markov (Carlin et al., 1992; Shephard & Pitt, 1997; Tanizaki, 2001).

O filtro de Kalman com parâmetros estimados

Quando os erros e o estado inicial X_1 são gaussianos, os estimadores do filtro de Kalman são os melhores estimadores não enviesados no sentido do erro quadrático médio mínimo. No entanto, as propriedades ótimas só podem ser garantidas quando todos os parâmetros do modelo de espaço de estado são conhecidos (Harvey, 1996).

A abordagem mais comum para lidar com modelos de espaço de estados com parâmetros estimados foca-se na análise do erro quadrático médio dos estimadores do filtro de Kalman. Quando o vetor de parâmetros desconhecidos θ é substituído por uma sua estimativa (por exemplo $\hat{\theta}_{MV}$), o erro quadrático médio dos estimadores é subestimado, uma vez que, por exemplo o erro de previsão vem dado por

$$P_{t|t-1}(\hat{\Theta}) = P_{t|t-1} + E \left[(X_{t|t-1} - X_{t|t-1}(\hat{\Theta}))^2 \right]. \quad (11)$$

Nesta perspetiva, a opção recai sobre a estimação da segunda parcela da Eq.10 através de metodologias baseadas no Bootstrap (Pfeffermann & Tiller, 2005; Bandyopadhyay & Lahiri, 2010, Rodríguez & Ruiz, 2012).

Recentemente, Costa & Monteiro (2016) apresentaram alguns resultados sobre a propagação do viés dos estimadores do filtro de Kalman, em particular, deduziram expressões não-recursivas para esse viés. Estes resultados permitem corrigir as estimativas iniciais dos parâmetros através de um procedimento iterativo de correção e, por conseguinte, as estimativas do filtro de Kalman, enquanto que na abordagem anterior corrige-se o erro quadrático médio dos estimadores do filtro de Kalman mantendo as suas estimativas pontuais.

4. Exemplo de aplicação

Considere-se um tanque cujos caudais de fluxo de entrada/saída se desconhecem com rigor, como se representa na Figura 2.

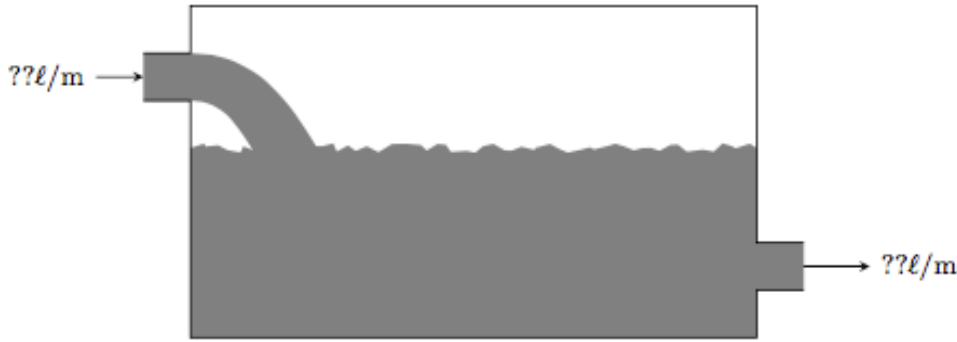


Figura 2: tanque com caudais de fluxo de entrada/saída desconhecidos.

Admita-se que o nível do líquido do tanque é monitorizado através de um instrumento que pode ser mecânico ou eletrónico, e que dispomos de 200 medições obtidas consecutivamente, em períodos de 1 minuto.

Seja Y_t a variável aleatória que representa a mediação do nível do tanque no minuto t obtido pelo instrumento utilizado e X_t o verdadeiro nível do tanque no mesmo instante. Assume-se que esta a medição e o verdadeiro nível do líquido do tanque se relacionam pelas equações

$$Y_t = X_t + e_t$$

$$X_t = \mu + \phi(X_{t-1} - \mu) + \varepsilon_t$$

onde $e_t \sim N(0, \sigma_e^2)$ e $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ são ruídos brancos não correlacionados. Este modelo contempla a hipótese de que o estado X_t seja um processo autorregressivo de ordem 1 estacionário, AR(1), caso $|\phi| < 1$.

Na Tabela 1 apresentam-se as estimativas de máxima verosimilhança, os respetivos erros padrão e os intervalos de confiança a 95% para os parâmetros do modelo.

Tabela 1: Estimativas de máxima verosimilhança e intervalos de confiança a 95% para os parâmetros do modelo.

parâmetro	estimativa	erro padrão	$I_{95\%}(\theta)$
μ	999,9145	0,0074	[999,75 ; 1000,08]
ϕ	0,5443	0,0151	[0,30 ; 0,79]
σ_e^2	0,5611	0,0167	[0,31 ; 0,81]
σ_ε^2	0,3412	0,0190	[0,07 ; 0,61]
$\ln L$	-281,0499		

As estimativas obtidas indicam que o nível do líquido do tanque é um processo estacionário uma vez que $|\hat{\phi}| = |0,5443| < 1$, com variância estimada igual a $\hat{\sigma}_{X_t}^2 = \hat{\sigma}_e^2 / (1 - \hat{\phi}^2) = 0,4848$. Assim, o ajustamento do modelo indica que a variância estimada do erro que podemos associar ao instrumento, $\hat{\sigma}_e^2 = 0,5611$, é superior à variância estimada inerente ao processo que representa o verdadeiro nível do líquido do tanque.

O ajustamento do modelo permitiu estimar o nível médio do líquido do tanque em 999,9145 litros, com um intervalo de confiança a 95% associado igual a [999,75;1000,08].

Com a aplicação do filtro de Kalman é possível obterem-se as estimativas filtradas do filtro de Kalman do nível do líquido do tanque, $X_{t|t}$, em cada instante e, se necessário, os respetivos intervalos de confiança. Na Figura 3 estão representadas as medições observadas do nível do líquido do tanque e as estimativas atualizadas (filtradas) obtidas pelo filtro de Kalman.

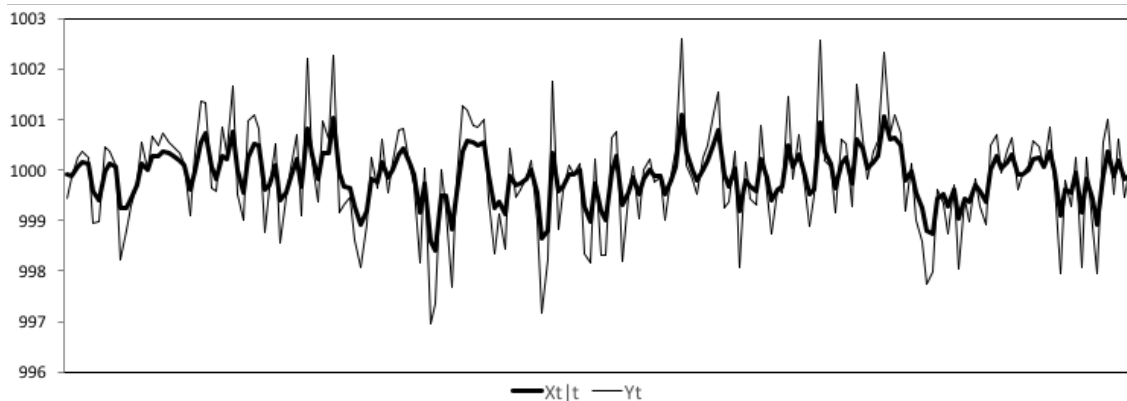


Figura 3: representação das observações no nível do líquido do tanque, Y_t , e as estimativas atualizadas (filtradas) obtidas pelo filtro de Kalman.

Desta forma, é possível fazer-se uma monitorização mais precisa dos caudais de fluxo de entrada/saída do do tanque tendo em consideração várias fontes de variabilidade e obtendo estimativas pontuais ou intervalares, se oportunas.

Referências

Alpuim, T. (1999). Noise variance estimators in state space models based on the method of moments, *Annales de l'ISUP*, **43** (2-3), 3-23.

- Bandyopadhyay, S., Lahiri, S.N. (2010). Resampling-based bias-corrected time series prediction. *Journal of Statistical Planning and Inference*, **140**, 3775–3788.
- Carlin, B. P., Polson, N. G., Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state space modeling. *Journal of the American Statistical Association*, **87**:493–500.
- Costa, M., Alpuim, T. (2010). Parameter estimation of state space models for univariate observations. *Journal of Statistical Planning and Inference*, **140** (7), 1889-1902.
- Costa, M., Monteiro, M. (2016). Bias-correction of Kalman filter estimators associated to a linear state space model with estimated parameters, *Journal of Statistical Planning and Inference*, **176**: 22 - 32.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1–38.
- Gonçalves, AM, Costa, M. (2013). Predicting seasonal and hydro-meteorological impact in environmental variables modelling via Kalman filtering, *Stochastic Environmental Research and Risk Assessment*, **27**(5), 1021-1038.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press, New Jersey.
- Harvey, A.C. (1996). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME--Journal of Basic Engineering*, 82 Series D, 35-45.
- Petris, G. (2010). An R Package for Dynamic Linear Models. *Journal of Statistical Software*, **36**(12), 1-16.
- Pfeffermann, D., Tiller, R. (2005). Bootstrap approximation to prediction MSE for state space-models with estimated parameters. *Journal of Time Series Analysis*, **21**, 219–236.
- Rodríguez, A., Ruiz, E. (2012). Bootstrap prediction mean squared errors of unobserved states based on the Kalman filter with estimated parameters. *Computational Statistics and Data Analysis*, **56**, 62–74.
- Shephard, N., Pitt, M. K. (1997). Likelihood Analysis of Non-Gaussian Measurement Time Series, *Biometrika*, **84**(3), 653-667.
- Tanizaki, H. (2001). Estimation of unknown parameters in non- linear and non-Gaussian state-space models. *Journal of Statistical Planning and Inference*, **96**, 301–323.